University of Saskatchewan
Department of Computational Science
Winter 1991

Cmpt 115 Final Examination

Closed Book
Time Limit: 3 hours

McCrosky & McCalla
26 April 1991

Instructions

You have three hours to write this exam. There are 100 marks, distributed as indicated next to each question. Write all answers directly on the examination paper. If insufficient space is available, indicate clearly that the answer is continued on the back of the current page, and continue on the backside. Do not become enchanted by any one question and fail to leave sufficient time for the others. In discussion questions you may use point-form if you like. Write legibly – no marks will be given for answers that cannot be read. Good luck!

Your name:

Your student number:

Your professor's name:

1) [2 marks] In natural language understanding, pragmatics refers to:

a) the structure of language
b) the context of language
c) the meaning of language
d) a and b
e) b and c
f) a and c
g) all three of a, b, and c
h) none of a, b, or c

2) [2 marks] Church's thesis (often called the Church-Turing thesis):

a) states that there are true but unprovable results
b) is used to prove that the halting problem is not computable
c) states that there are no more powerful ways of stating computation than a Turing machine
d) states that Turing machines are equivalent to type-0 grammars
e) none of the above

3) [2 marks] A heuristic:

a) cannot be implemented as an algorithm
b) is an *ad hoc* "rule of thumb"
c) is an essential part of any search technique
d) a and b
e) b and c
f) a and c
g) all three of a, b, and c
h) none of a, b, or c

4) [2 marks] A last-in-first-out queuing discipline:

a) is used by a stack
b) is equivalent to first-in-last-out
c) can be implemented using linked or non-linked structures
d) a and b
e) a and c
f) b and c
g) all of a, b, and c
h) none of a, b, or c

5) [2 marks] In sequential file processing if the master file has m records, the transaction file has t records, and m is very much greater than t, then the merge-update loop has the following computational complexity:

    a) $O(m + t)$
    b) $O(m * t)$
    c) $O(m \log t)$
    d) $O(t \log m)$
    e) $O(m)$
    f) $O(t)$
    g)   none of the above

6) [3 marks] In computer game playing, why is it necessary to do lookahead search? In a given board position, why not just apply an evaluation function to all possible successor board positions, then make the move that leads to the successor board with the highest evaluation?

7) [3 marks] Define the following:

    a)   halting problem

   b)   context sensitive grammar

    c)   top-down parsing

8) [4 marks] Indicate for each of the following applications whether direct or sequential file processing is more appropriate:

a)   a banking machine

b)   writing payroll cheques for a large corporation

c)   an office information system to be used by managers as an aid to decision making

d)   a cash register that retrieves prices from laser scanned universal product codes (as at Safeway and SuperValu(e)).

9) [5 marks] In a doubly linked list, each element in the list has pointers to both its predecessor element and its successor element. In the following, assume there is a list of numbers:

$$1 \ 2 \ 7 \ 9 \ 24 \ 97$$

a)   Draw a doubly linked version of this list.

b)   Show Pascal declarations (declarations only) for a doubly linked version of the list.

c)   What advantages (relative to insertion, deletion, and search) would a doubly linked list have over a singly linked list? What disadvantages?

10) [5 marks] Assume you are figuring out how much wall to wall carpet to buy to cover the floor of an awkwardly shaped room. You have made measurements of all the nooks and crannies using a tape measure marked in inches. You must now determine how many square yards of carpet to buy. If you were using a computer to make this calculation, would you convert all the linear inches measurements into yards and then compute the area to be covered, or would you calculate the total area to be covered in square inches and then convert to square yards? Argue your case in terms of truncation and round-off error analysis principles.

11) [8 marks] Give a grammar that recognizes the same language as the finite state machine diagrammed below:
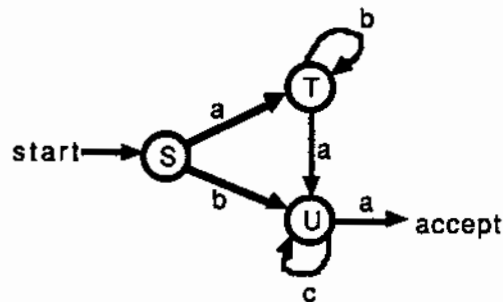
What type of grammar is this (what is its Chomskian classification)?

12) [7 marks] Consider the following table of symbols and their frequency of occurrence in a given text:

|			
d	0.37	c	0.09
r	0.25	k	0.08
e	0.14	y	0.07

a) Generate the Huffman encoding tree corresponding to these frequencies.

b) What is the encoding of the string "creedy"?

13) **[10 marks]** Assuming availability of an algorithm for integration and one for finding roots (zero crossings) of equations, describe a high level algorithm to compute the area above $f(x)$ and below $g(x)$ shaded in the following diagram. (Our interest is in one interval where $g(x)$ is consistently greater than $f(x)$. Assume $f(x)$ is consistently greater than $g(x)$ everywhere else.)
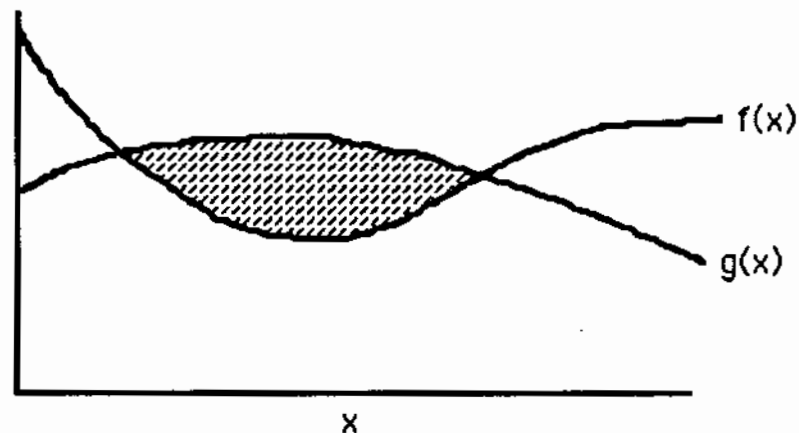
14) **[10 marks]** A *reachability matrix* for an undirected graph is an *n*-by-*n* matrix, where *n* is the number of vertices in the graph. Entries in the matrix, *rm[i,j]*, are one if there exists a path between the ith and j^{th} nodes (or if $i = j$), and zero otherwise. The diagram below gives a graph and its reachability matrix.

	1	2	3	4	5	6
1	1	0	0	1	0	0
2	0	1	0	0	0	0
3	0	0	1	0	1	1
4	1	0	0	1	0	0
5	0	0	1	0	1	1
6	0	0	1	0	1	1
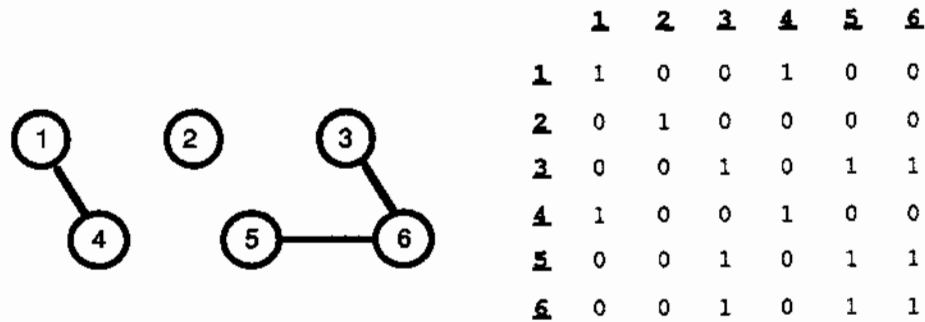
Given a reachability matrix, devise a high level algorithm to determine the <u>number</u> of disconnected subgraphs present in the graph (e.g., there are three disconnected subgraphs in the graph above, {1, 4}, {2}, and {3, 5, 6})

15) [10 marks] Consider the following game tree. The square nodes are maximizing nodes, the circle nodes are minimizing nodes, and the numbers inside the leaf nodes are the result of the application of an evaluation function at that point.
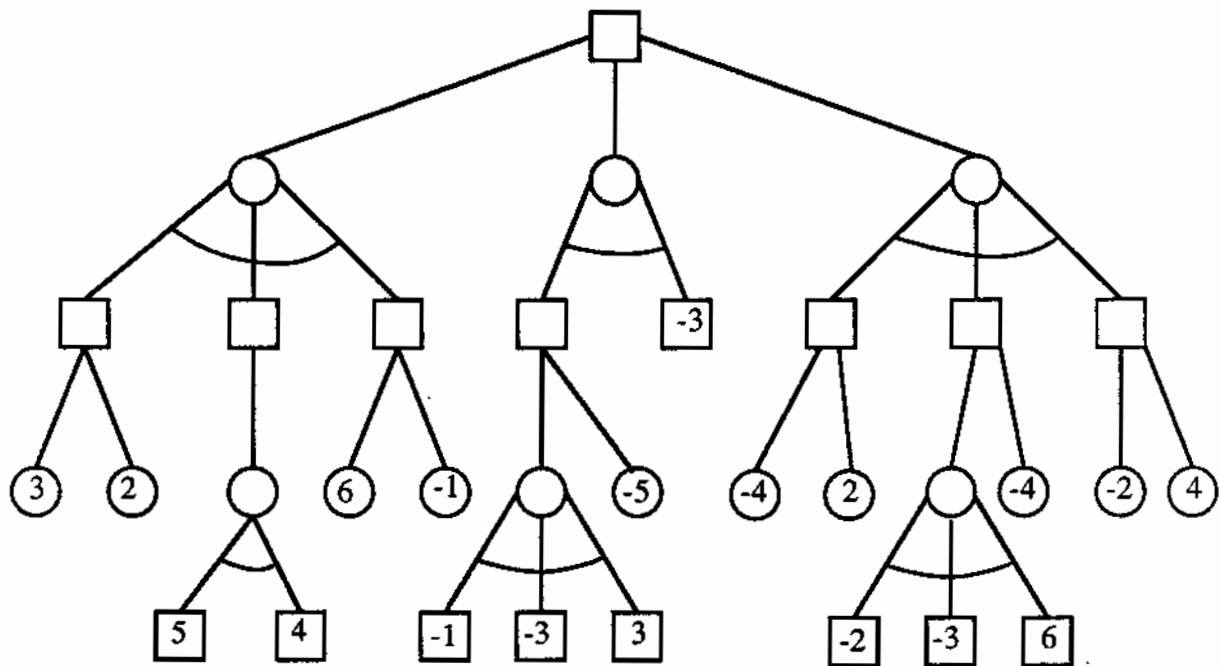
Show what branches of this tree will be cutoff by alpha-beta pruning, assuming the tree is generated left to right, depth first. Indicate next to each cutoff whether it is an alpha cutoff or a beta cutoff. Is the final back-up value the same after alpha-beta pruning as it would have been without this pruning?

16) [10 marks] When a customer is making a reservation on an airline reservation system a lot of information must be kept. Customers must be booked onto particular flights on particular days; the type of ticket and its price must be recorded; customer name, address and phone number should be kept; sources and destinations of the flights must be recorded, as must departure and arrival times; flight numbers and airlines must be kept. Assume you are charged with the responsibility of designing an airline reservation system and you have decided to store this information in a set of relations in a relational database.

   a) Indicate the various relations that you will create, including the attributes in each relation. Note that it is important, if at all possible, to eliminate redundancy so that the same information is not recorded more than once. Give an example tuple for each relation.

b)   Using the relational database query operations select, join, and project formulate queries for the following questions that could be posed of your database:

"What flights are Mr. Smith booked on?"

"What are the names and home addresses of all people booked on full fare economy flights on Air Canada from Winnipeg to Saskatoon in May 1991?"

17) [15 marks] Have you ever wanted to have a personalized license plate that read "I HATE EXAMS". Well here's your chance. The Devine government, as a pre-election goody, is offering arbitrarily long character strings on license plates. (Aren't you glad you didn't vote NDP?).

As with most government programs, there are complications. A devine province clearly cannot afford to have foul language printed on its citizens' license places. And of course there cannot be duplicates. So the government is planning to search for all requested character strings in a database of all possible strings.

This database is stored as a tree of all possible character strings, as shown below (Well, maybe not really all strings of all lengths, as that would be infinite. But for the purposes of this question assume the tree is infinite.) The root node contains the empty character string. Each node has 27 children: one for each letter of the alphabet and one for the blank character. Each such child node contains the string of its parent node with its particular character concatenated at the end. For example, in the diagram below, the node "ab" is a child of the node "a" with "b" concatenated.
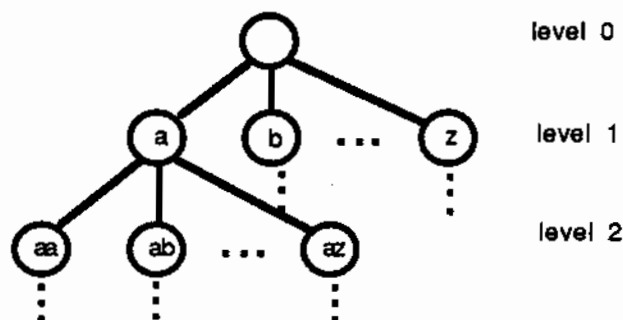
Associated with each node are two boolean flags: one indicates whether the phrase is foul, the other whether it has been used by a previous license plate.

If one did a depth-first search of this tree, the program would recur forever down the left-most branch of the tree, and never fully search the tree.

Instead, you are to solve the problem using a "breadth-first" search. You are to write a program to search the tree in order by "levels", where the levels are marked in the diagram above. A high-level algorithm for breadth-first searching the tree described above is given below:

a) put a pointer to the root of the tree in a queue as the first element.
b) remove a node from the head of the queue.
c) if the node contains a string longer than the string you are checking, quit and write "not found" (why does this work?).
d) if the node matches the string, quit and write "found".
e) else put the children of the node in the queue
f) go to (b).

Write a Pascal program to breadth-first search the tree for an arbitrary license plate string. You may assume queue insert and delete procedures exist.